RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/921,922A

DATE: 01/04/2002
TIME: 15:05:42

Input Set : A:\16516-107 seq listing.txt
Output Set: N:\CRF3\01042002\I921922A.raw

ENTERED

```
 1 <110> APPLICANT: Boronat, Albert;
 2         Campos, Narciso;
 3         Rodriguez-Concepcion, Manuel;
 4         Rohmer, Michel;
 5         Seeman, Myriam;
 6         Valentin, Henry E.;
 7         Venkatesh, Tyamagondlu V.;
 8         Venkatramesh, Mylavarapu
10 <120> TITLE OF INVENTION: Methyl-D-Erythritol Phosphate Pathway Genes
12 <130> FILE REFERENCE: 16516.107/35-21(51897)US
14 <140> CURRENT APPLICATION NUMBER: US/09/921,922A
14 <141> CURRENT FILING DATE: 2001-12-19
14 <150> PRIOR APPLICATION NUMBER: US 60/223,483
15 <151> PRIOR FILING DATE: 2000-08-07
17 <160> NUMBER OF SEQ ID NOS: 85
19 <210> SEQ ID NO: 1
20 <211> LENGTH: 2520
21 <212> TYPE: DNA
22 <213> ORGANISM: Arabidopsis thaliana
24 <220> FEATURE:
25 <221> NAME/KEY: CDS
26 <222> LOCATION: (154)..(2376)
28 <400> SEQUENCE: 1
30 aaaaatcgtc aatccctctc aaactcttct caccactaat ttcttcctct ggaacattct      60
32 cttctctatt attttgattc ccttggcctc aacactggtt tctcaattgc atgatcttgg     120
34 ctcgtcttca gttactttga ttcactgaga aaa atg gcg act gga gta ttg cca      174
35                                     Met Ala Thr Gly Val Leu Pro
36                                      1               5
38 gct ccg gtt tct ggg atc aag ata ccg gat tcg aaa gtc ggg ttt ggt      222
39 Ala Pro Val Ser Gly Ile Lys Ile Pro Asp Ser Lys Val Gly Phe Gly
40         10              15              20
42 aaa agc atg aat ctt gtg aga att tgt gat gtt agg agt cta aga tct      270
43 Lys Ser Met Asn Leu Val Arg Ile Cys Asp Val Arg Ser Leu Arg Ser
44     25              30              35
46 gct agg aga aga gtt tcg gtt atc cgg aat tca aac caa ggc tct gat      318
47 Ala Arg Arg Arg Val Ser Val Ile Arg Asn Ser Asn Gln Gly Ser Asp
48 40              45              50              55
50 tta gct gag ctt caa cct gca tcc gaa gga agc cct ctc tta gtg cca      366
51 Leu Ala Glu Leu Gln Pro Ala Ser Glu Gly Ser Pro Leu Leu Val Pro
52             60              65              70
54 aga cag aaa tat tgt gaa tca ttg cat aag acg gtg aga agg aag act      414
55 Arg Gln Lys Tyr Cys Glu Ser Leu His Lys Thr Val Arg Arg Lys Thr
56             75              80              85
58 cgt act gtt atg gtt gga aat gtc gcc ctt gga agc gaa cat ccg ata      462
59 Arg Thr Val Met Val Gly Asn Val Ala Leu Gly Ser Glu His Pro Ile
60         90              95              100
62 agg att caa acg atg act act tcg gat aca aaa gat att act gga act      510
```

RAW SEQUENCE LISTING   DATE: 01/04/2002
PATENT APPLICATION: US/09/921,922A   TIME: 15:05:42

Input Set : A:\16516-107 seq listing.txt
Output Set: N:\CRF3\01042002\I921922A.raw

```
63 Arg Ile Gln Thr Met Thr Thr Ser Asp Thr Lys Asp Ile Thr Gly Thr
64     105             110             115
66 gtt gat gag gtt atg aga ata gcg gat aaa gga gct gat att gta agg      558
67 Val Asp Glu Val Met Arg Ile Ala Asp Lys Gly Ala Asp Ile Val Arg
68 120             125             130             135
70 ata act gtt caa ggg aag aaa gag gcg gat gcg tgc ttt gaa ata aaa      606
71 Ile Thr Val Gln Gly Lys Lys Glu Ala Asp Ala Cys Phe Glu Ile Lys
72             140             145             150
74 gat aaa ctc gtt cag ctt aat tac aat ata ccg ctg gtt gca gat att      654
75 Asp Lys Leu Val Gln Leu Asn Tyr Asn Ile Pro Leu Val Ala Asp Ile
76             155             160             165
78 cat ttt gcc cct act gta gcc tta cga gtc gct gaa tgc ttt gac aag      702
79 His Phe Ala Pro Thr Val Ala Leu Arg Val Ala Glu Cys Phe Asp Lys
80             170             175             180
82 atc cgt gtc aac cca gga aat ttt gcg gac agg cgg gcc cag ttt gag      750
83 Ile Arg Val Asn Pro Gly Asn Phe Ala Asp Arg Arg Ala Gln Phe Glu
84     185             190             195
86 acg ata gat tat aca gaa gat gaa tat cag aaa gaa ctc cag cat atc      798
87 Thr Ile Asp Tyr Thr Glu Asp Glu Tyr Gln Lys Glu Leu Gln His Ile
88 200             205             210             215
90 gag cag gtc ttc act cct ttg gtt gag aaa tgc aaa aag tac ggg aga      846
91 Glu Gln Val Phe Thr Pro Leu Val Glu Lys Cys Lys Lys Tyr Gly Arg
92             220             225             230
94 gca atg cgt att ggg aca aat cat gga agt ctt tct gac cgt atc atg      894
95 Ala Met Arg Ile Gly Thr Asn His Gly Ser Leu Ser Asp Arg Ile Met
96             235             240             245
98 agc tat tac ggg gat tct ccc cga gga atg gtt gaa tct gcg ttt gag      942
99 Ser Tyr Tyr Gly Asp Ser Pro Arg Gly Met Val Glu Ser Ala Phe Glu
100         250             255             260
102 ttt gca aga ata tgt cgg aaa tta gac tat cac aac ttt gtt ttc tca     990
103 Phe Ala Arg Ile Cys Arg Lys Leu Asp Tyr His Asn Phe Val Phe Ser
104     265             270             275
106 atg aaa gcg agc aac cca gtg atc atg gtc cag gcg tac cgt tta ctt     1038
107 Met Lys Ala Ser Asn Pro Val Ile Met Val Gln Ala Tyr Arg Leu Leu
108 280             285             290             295
110 gtg gct gag atg tat gtt cat gga tgg gat tat cct ttg cat ttg gga     1086
111 Val Ala Glu Met Tyr Val His Gly Trp Asp Tyr Pro Leu His Leu Gly
112             300             305             310
114 gtt act gag gca gga gaa ggc gaa gat gga cgg atg aaa tct gcg att     1134
115 Val Thr Glu Ala Gly Glu Gly Glu Asp Gly Arg Met Lys Ser Ala Ile
116             315             320             325
118 gga att ggg acg ctt ctt cag gac ggg ctc ggt gac aca ata aga gtt     1182
119 Gly Ile Gly Thr Leu Leu Gln Asp Gly Leu Gly Asp Thr Ile Arg Val
120         330             335             340
122 tca ctg acg gag cca cca gaa gag gag ata gat ccc tgc agg cga ttg     1230
123 Ser Leu Thr Glu Pro Pro Glu Glu Glu Ile Asp Pro Cys Arg Arg Leu
124     345             350             355
126 gct aac ctc ggg aca aaa gct gcc aaa ctt caa caa ggc gca ccg ttt     1278
127 Ala Asn Leu Gly Thr Lys Ala Ala Lys Leu Gln Gln Gly Ala Pro Phe
```

RAW SEQUENCE LISTING   DATE: 01/04/2002
PATENT APPLICATION: US/09/921,922A   TIME: 15:05:42

Input Set : A:\16516-107 seq listing.txt
Output Set: N:\CRF3\01042002\I921922A.raw

```
128 360                    365                   370                        375
130 gaa gaa aag cat agg cat tac ttt gat ttt cag cgt cgg acg ggt gat      1326
131 Glu Glu Lys His Arg His Tyr Phe Asp Phe Gln Arg Arg Thr Gly Asp
132                    380                   385                   390
134 cta cct gta caa aaa gag gga gaa gag gtt gat tac aga aat gtc ctt      1374
135 Leu Pro Val Gln Lys Glu Gly Glu Glu Val Asp Tyr Arg Asn Val Leu
136                    395                   400                   405
138 cac cgt gat ggt tct gtt ctg atg tcg att tct ctg gat caa cta aag      1422
139 His Arg Asp Gly Ser Val Leu Met Ser Ile Ser Leu Asp Gln Leu Lys
140                410                   415                   420
142 gca cct gaa ctc ctc tac aga tca ctc gct aca aag ctt gtc gtg ggt      1470
143 Ala Pro Glu Leu Leu Tyr Arg Ser Leu Ala Thr Lys Leu Val Val Gly
144        425                   430                   435
146 atg cca ttc aag gat ctg gca act gtt gat tca atc tta tta aga gag      1518
147 Met Pro Phe Lys Asp Leu Ala Thr Val Asp Ser Ile Leu Leu Arg Glu
148 440                   445                   450                   455
150 cta ccg cct gta gat gat caa gtg gct cgt ttg gct cta aaa cgg ttg      1566
151 Leu Pro Pro Val Asp Asp Gln Val Ala Arg Leu Ala Leu Lys Arg Leu
152                    460                   465                   470
154 att gat gtc agt atg gga gtt ata gca cct tta tca gag caa cta aca      1614
155 Ile Asp Val Ser Met Gly Val Ile Ala Pro Leu Ser Glu Gln Leu Thr
156                475                   480                   485
158 aag cca ttg ccc aat gcc atg gtt ctt gtc aac ctc aag gaa cta tct      1662
159 Lys Pro Leu Pro Asn Ala Met Val Leu Val Asn Leu Lys Glu Leu Ser
160                490                   495                   500
162 ggt ggc gct tac aag ctt ctc cct gaa ggt aca cgc ttg gtt gtc tct      1710
163 Gly Gly Ala Tyr Lys Leu Leu Pro Glu Gly Thr Arg Leu Val Val Ser
164        505                   510                   515
166 cta cga ggc gat gag cct tac gag gag ctt gaa ata ctc aaa aac att      1758
167 Leu Arg Gly Asp Glu Pro Tyr Glu Glu Leu Glu Ile Leu Lys Asn Ile
168 520                   525                   530                   535
170 gat gct act atg att ctc cat gat gta cct ttc act gaa gac aaa gtt      1806
171 Asp Ala Thr Met Ile Leu His Asp Val Pro Phe Thr Glu Asp Lys Val
172                540                   545                   550
174 agc aga gta cat gca gct cgg agg cta ttc gag ttc tta tcc gag aat      1854
175 Ser Arg Val His Ala Ala Arg Arg Leu Phe Glu Phe Leu Ser Glu Asn
176        555                   560                   565
178 tca gtt aac ttt cct gtt att cat cac ata aac ttc cca acc gga atc      1902
179 Ser Val Asn Phe Pro Val Ile His His Ile Asn Phe Pro Thr Gly Ile
180        570                   575                   580
182 cac aga gac gaa ttg gtg att cat gca ggg aca tat gct gga ggc ctt      1950
183 His Arg Asp Glu Leu Val Ile His Ala Gly Thr Tyr Ala Gly Gly Leu
184        585                   590                   595
186 ctt gtg gat gga cta ggt gat ggc gta atg ctc gaa gca cct gac caa      1998
187 Leu Val Asp Gly Leu Gly Asp Gly Val Met Leu Glu Ala Pro Asp Gln
188 600                   605                   610                   615
190 gat ttt gat ttt ctt agg aat act tcc ttc aac tta tta caa gga tgc      2046
191 Asp Phe Asp Phe Leu Arg Asn Thr Ser Phe Asn Leu Leu Gln Gly Cys
192                620                   625                   630
```

RAW SEQUENCE LISTING   DATE: 01/04/2002
PATENT APPLICATION: US/09/921,922A   TIME: 15:05:42

Input Set : A:\16516-107 seq listing.txt
Output Set: N:\CRF3\01042002\I921922A.raw

```
194 aga atg cgt aac act aag acg gaa tat gta tcg tgc ccg tct tgt gga   2094
195 Arg Met Arg Asn Thr Lys Thr Glu Tyr Val Ser Cys Pro Ser Cys Gly
196         635             640             645
198 aga acg ctt ttc gac ttg caa gaa atc agc gcc gag atc cga gaa aag   2142
199 Arg Thr Leu Phe Asp Leu Gln Glu Ile Ser Ala Glu Ile Arg Glu Lys
200         650             655             660
202 act tcc cat tta cct ggc gtt tcg atc gca atc atg gga tgc att gtg   2190
203 Thr Ser His Leu Pro Gly Val Ser Ile Ala Ile Met Gly Cys Ile Val
204     665             670             675
206 aat gga cca gga gaa atg gca gat gct gat ttc gga tat gta ggt ggt   2238
207 Asn Gly Pro Gly Glu Met Ala Asp Ala Asp Phe Gly Tyr Val Gly Gly
208 680             685             690             695
210 tct ccc gga aaa atc gac ctt tat gtc gga aag acg gtg gtg aag cgt   2286
211 Ser Pro Gly Lys Ile Asp Leu Tyr Val Gly Lys Thr Val Val Lys Arg
212             700             705             710
214 ggg ata gct atg acg gag gca aca gat gct ctg atc ggt ctg atc aaa   2334
215 Gly Ile Ala Met Thr Glu Ala Thr Asp Ala Leu Ile Gly Leu Ile Lys
216             715             720             725
218 gaa cat ggt cgt tgg gtc gac ccg ccc gtg gct gat gag tag           2376
219 Glu His Gly Arg Trp Val Asp Pro Pro Val Ala Asp Glu
220         730             735             740
222 atttcaaaac ggagaaagat gggtgggcca ttctttgaaa actgtgagag aagatatata   2436
224 tatttgtgtg tgtatatcat ctgtttgttg tgtattgcat catcattttg aacaaatgtc   2496
226 caaatctctt aagttgataa aagt                                         2520
229 <210> SEQ ID NO: 2
230 <211> LENGTH: 33675
231 <212> TYPE: DNA
232 <213> ORGANISM: Oryza sativa
234 <220> FEATURE:
235 <221> NAME/KEY: CDS
236 <222> LOCATION: (6924)..(7019),(7163)..(7269),(7344)..(7444),(7525)..(7634),
237 <222> LOCATION: (7694)..(7813),(7923)..(8153),(8253)..(8369),(8515)..(8589),
238 <222> LOCATION: (9012)..(9071),(9163)..(9225),(9328)..(9472),(9589)..(9730),
239 <222> LOCATION: (9951)..(10028),(10134)..(10293),(10694)..(10798),
240 <222> LOCATION: (11028)..(11129)
242 <220> FEATURE:
243 <221> NAME/KEY: unsure
244 <222> LOCATION: (1..33675)
245 <223> OTHER INFORMATION: unsure at all n locations
247 <400> SEQUENCE: 2
249 cttaaccctc gccgactgcc tggagattcg tgccgatcga tacacgtggc agcgcctaac     60
251 gcgtaacccc tccctcactt ggagattcgt gcaagcaact cgattaatgc attaatgctg    120
253 tcgcgtaggt ttccctacgg aagagctgag tttcgtaacg aaaaaaaccg gccacgtttc    180
255 gcatcgagcc tactttaatt agcgtgggaa aataattcaa agtagcgacc tgtaccctgt    240
257 ggcaacctag cgcgcgcggc catggctctt gttccgctcg tgacagtgct cctgttcgcc    300
259 ggctcatgcc tcggatcagc gccgccgacg acatcgccgg cggcgtcggc ggcgtccacg    360
261 gcgacacgta cggtagtagt cgacggcatt acggccatct acaaacctcg gcgactcgct    420
263 gtcggacacc gcaacctcgc caggcaaggc gccaccggcg ggctgctccg gtacaccacg    480
265 aggcttccct acggcgtcac cgtcggccgc gccaccggcc ggtgctccga cggctacctc    540
```

```
267 atcatcgact tcctcggtga cgtcatcagt ttaatttctc tctctcttcc gtctgaaaaa   600
269 tggaagaaac aatattatat tacgttatat atatatgcgt ttttgtttcg gattaaattg   660
271 tggatatgat cgatcgatgt gcagctagag atcttggcct ccctctgctc aacccgtacc   720
273 tcgacgaggg cgcggacttc gcccacggcg tcaacttcgc cgtcgccggc gccaccgcgc   780
275 tcaacacgac ggcgctcgcc gccaggcgga tcaccgtccc ccacaccaac agcccctcg   840
277 acgtgcagct cagatttttt ttgtttaga gaagggtatt ttttacccgg cctctacatc   900
279 caaccggata tatacggcta ttgaagtagg gaacttaacc ctgtaaacaa tccatccata   960
281 gaggatatga acctaagacc ttgaggtact acttcaaccg gatatatacg tgcagctcag  1020
283 atggttcaag gaattcatga actccacaac tagttctcct caaggtgaac gaacaaactg  1080
285 aaacgcattt cagcttaatt tcgaccggtg cctgatcagt gccagtcagc aatgctgtat  1140
287 ctcacaaata attaagctaa tgtacagctt ttcagtgcta gaatgacttt catatagaga  1200
289 aatcttgtgt tatatatata tacttttttc tgaaagaaaa aagttctttt gtgtgagcat  1260
291 tgcattgcag agatccgtga aaagctgtcg aagtcactgg ttatgctggg agagatcgga  1320
293 ggaaacgact acaactacgc cttcctccag acctggccga tggacggtgg atacagcctc  1380
295 ggcaacgtca cacgcatgat cgaaagcgtt gccacggccg tcgatcttgt accggaagtc  1440
297 gtgcagtcca tagccagcgc agccaaggta cacaccattc ttttccatta atttttggga  1500
299 ccttattttt aaaataataa tcctggctac aaagtaatta attaagaact aaattaattt  1560
301 ttgtgggttt tgtgacacag gaggtgctcg acatgggcgc gacgcgggtg gtgatcccgg  1620
303 gcaacctccc gctgggttgc gtgccgagct acatgagcgc ggtgaacgcg acggaccggg  1680
305 cggcgtacga cgcccgcgga tgcctcgtcg cgctcaacct cttcgcggcg ctgcacaacg  1740
307 cgtggctgcg ccgcgccgtc ggggagctgc ggcgcgcgta ccggggcgcc gcggtggtcg  1800
309 cgtacgcgga ctactccgcc gcgtacgccg cgacgctgga cggggcagcg gcgctcggct  1860
311 tcgacgagcg gcgcgtgttc agggcgtgct gcggcaaggg cggcgggggc gcgtacgggt  1920
313 tcgacgtgcg cgcgatgtgc ggcgcgccgg ggacggcggc gtgcgcggac ccggggaggt  1980
315 acgtgagctg ggacggcgtc cacctgacgc agcgcgcgta cggcgtcatg gccgagctgc  2040
317 tgttccgccg tggcctcgtg cacccgcctc cgataaattt cacgaacagc gcgcgcgcgt  2100
319 gaggcggtgt tgcatggctt gcgcgttttt tctgatcaaa actactcaag tttgagccgt  2160
321 tttgatttat aaataaaacc atatgcgatt ttgctaaacg tttgtcgcgt gatttctctt  2220
323 cggaagaaaa aatctcaccc gagtgatgca taggcggtcc caaccatatg tgccctgacc  2280
325 tttctctgct tccttcgcgt cgtgcactga caacctcaca gtatgttttt ggtatgggcg  2340
327 cttgcggccc aactcaatct gtaatacatt gggctgtcgt attgggtttg ttggacttca  2400
329 tagactggat cggagaaagt tgggtaattg acttttttcat ttttgctata aaatgattaa  2460
331 ttaaacagtc taggataatt actgtagact ctaataatat tgtttggtta agtattatta  2520
333 tacattcctg tatttgacac tctaagagca tggccaagag ttgcctgaaa gtctcttcct  2580
335 aaatctgcct ttcattctct aatgagaatt taaggattaa aaatatactt attttcaata  2640
337 gacagcataa atttaattcc ctagaataaa aaaatgcccc cctaacaaca gaaattagat  2700
339 tcctctaccc gcacctcatc agatcgctcg atttaagatc acgccatctg acaccgccct  2760
341 cccgctcgct cttctctagt gtgggagtct cgcgctcaag agacggaaat cgggaacaag  2820
343 aatgattcct agcttagcga gaatgaaggg gaagacatat gtcatcctta cacccacata  2880
345 agtatgccct agcacaaggg atgaaaacgg atcgaaaacg gatggaaact agctttatca  2940
347 tattcgtttt catttttttt tcggaatcgg attcgaaatc gaaaactcgg atacggaaat  3000
349 aaaattgaat attatcgaat acagatacgg agcgaatata agatggaacg aatacagtag  3060
351 cgaatattta ccggtatata aaaaacccct caaattgagt ttcttgatta agaaagagat  3120
353 atcgcttatt attttagtta aatatctcca acatttatat cgtcaatttt atagacggtt  3180
355 ccacaatcgt atgtgaaaat cgattttcat ggttgttcct ctaagagatc catatgcaaa  3240
357 tatgattatc attttctatt ccaagacctt ttactagatg tataacttat ttaccattgc  3300
359 ataaattgga gatgttattt attttacttc acatcttcga aacttgtaat gtatgtatta  3360
361 tactttaaat gctttcaagt acaaatgtta taaactacaa agtggtagat cccgttgagc  3420
363 tctacaactt tgatatggaa cacatctcca tcagatgtcg tttgaattgt agatctgaga  3480
```

Use of n and/or Xaa has been detected in the Sequence Listing.
Review the Sequence Listing to insure a corresponding
explanation is presented in the <220> to <223> fields of
each sequence using n or Xaa.

VERIFICATION SUMMARY　　　　　　DATE: 01/04/2002
PATENT APPLICATION: US/09/921,922A　　TIME: 15:05:43

Input Set : **A:\16516-107 seq listing.txt**
Output Set: **N:\CRF3\01042002\I921922A.raw**

```
L:14 M:270 C: Current Application Number differs, Replaced Current Application No
L:14 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:821 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2
L:1759 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5
L:1763 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5
L:1765 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5
L:1767 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5
L:1769 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5
L:1798 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6
L:1881 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9
L:1883 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9
L:1885 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9
L:1887 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9
L:1889 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9
L:1948 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11
L:1950 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11
L:1952 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11
L:1993 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13
L:1995 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13
L:2064 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15
L:2217 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:2219 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:2310 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:23
L:2515 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:32
L:2606 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:35
L:2667 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:37
L:2834 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:43
L:2836 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:43
L:2840 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:43
```